Available online at www.sciencedirect.com

ScienceDirect

IERI Procedia 6 (2014) 63 – 71

Procedia IERI

www.elsevier.com/locate/procedia

2013 International Conference on Future Software Engineering and Multimedia Engineering

Soft Sensor for CO_x Content in Tail Gas of PX Oxidation Side Reactions Based on Particle Filters and EM Algorithm

Ranran Liu [a,b], Zhengming Li [a,*]

[a] *School of Electrical and Information Engineering, Jiangsu University,Zhenjiang 212013,China*
[b] *Changzhou Liu Guojun Vocational Technology College, Changzhou 213004,China*

Abstract

Based on expectation-maximization algorithm, parameter estimation was proposed for data-driven nonlinear models in this work. On this basis, particle filters were used to approximately calculate integrals, deriving EM algorithm based on particle filter. And the effectiveness of using the proposed algorithm for the soft sensor of CO_x content in tail gas of PX oxidation side reactions was verified through simulation results.

Keywords: Expectation-maximization algorithm; particle filtering; data-driven； soft sensor;

1. Introduction

We have been committed to improving the performance of process equipments through effective control in industrial process, achieving the enhancement in productivity and reduction in energy consumption. However, several key parameters can’t be directly measured through sensors due to technical or economic reasons in practical production, such as the concentration of product components in a distillation column, the cell

* *Li Zhengming.* Tel.: .+0-511-887-90618; fax: .+0-511-887-91245.
E-mail address: lzming@ujs.edu.cn

doi:10.1016/j.ieri.2014.03.011

concentration of a biological fermentation tank, the silicon content in BF (blast furnace) hot metal as well as the concentration of reactants, conversion rate and catalyst activity in a chemical reactor. One of the two traditional solutions is using on-line analyzer control. Although the method can achieve rapid detection, large investment in equipments, high maintenance costs and big hysteresis can influence the moderating effect. The other method is off-line experimental analysis to ensure the measurement accuracy of data. However, the long measuring time of the method was not conducive to real-time processing. Consequently, soft sensing technique came into being. Soft sensing technique selects several easily measured variables (or so-called auxiliary variables) to infer and estimate the important variables that are difficult to measure or temporarily can't be measured (or so-called dominant variable) through forming several mathematical relationships. The essence of the method is to establish the relation model between auxiliary and dominant variables (Brosillow 1978).

Soft sensor has been widely used (P. Kadleca, B. Gabrys and S. Strandtb 2009; S. Khatibisepehr and B. Huang 2008; W. Yan, H. Shao and X. Wang 2004). Based on conventional detection, it utilizes the relationship between auxiliary and dominant variables to obtain the estimates of dominant variables through calculation of software. There're many soft sensing methods for research and application, divided into three categories according to modeling approaches: mechanism model, grey-box model and data-driven model. Although there are a variety of structures of data-driven models, the steady-state and dynamic models are mostly studied. Considering the change of data over time, the commonly used dynamic models include state-space model and input-output model, such as autoregressive models (e.g., ARX, ARMAX), output-error (OE) models and Box-Jenkins (BJ) models (P.K. Pearson 2004). And the nonlinear dynamic model represents a nonlinear dynamic system in the above models. Thus the proposed soft sensing technique mainly adopts such models in this work. Recently, sampling based on SMC (Sequential Monte Carlo) (A. Doucet, N. de Freitas and N. Gordon 2001), a parameter estimation method combining particle filter and EM algorithm [8] was proposed. In Literature(R.B. Gopaluni 2008) EM algorithm, particle filter and smoother were used to approximate logarithmic likelihood functions. Using the pointwise state estimation, each density function after state smoothing will be calculated in every iteration of EM algorithm, which is a great deal of calculation. In the work, the expectation function was calculated by using particle filter. Replacing smoother with filter can not only reduce calculation but also bring acceptable estimated performance.

Here is the organization: Section 2 described the parameter estimation of nonlinear state-space model in the framework of EM algorithm. In Section 3, the proposed algorithm was applied to PX oxidation side reactions for estimating COx content in the tail gas. Section 4 discussed several problems, and Section 5 made a conclusion based on simulation results.

2. Parameter estimation of state-space model

2.1. Model structure

The nonlinear state-space model was given:

$$x_t = f(x_{t-1}, u_{t-1}, \theta) + \omega_t \tag{1}$$

$$y_t = h(x_t, \theta) + v_t \tag{2}$$

where θ is a vector of system parameter, and x_t , u_t and y_t are respectively state, measurement of input and output. ω_t and v_t are covariance: respectively the white noise of Q and R. Input $\{u_1, u_2 ......, u_t\}$ was known, omitted in the subsequent derivation for simplification. Let X represent hidden state $\{x_1, x_2 ......, x_T\}$,and Y_0 represents $\{y_1, y_2 ......, y_T\}$ to embody the corresponding measurement of output.

2.2. Expectation-maximization algorithm

In statistics, expectation-maximization algorithm searches for the maximum likelihood parameter estimation in the probability model that relies on unobservable latent variables. The EM algorithm is alternately calculated in two steps: The first step is calculating expectation (E), in which the existing estimates of latent variables were used to calculate their maximum likelihood estimates; The second step is maximization (M), where the maximum likelihood estimates from Step E were utilized to calculate the parameter values. And the estimates of parameters obtained in Step M were used in the next Step E, thus this process alternately continues.

For state-space model (1) and (2), Step E is to calculate the Q function, defined as follows:

$$Q(\theta \mid \theta^k) = E_{X \mid Y_0, \theta^k} \left\{ \log[p = (Y_0, X \mid \theta)] \right\} = \int \log[p = (Y_0, X \mid \theta)] p(X \mid Y_0, \theta^k) dX \tag{3}$$

Step M is to determine the value of θ when the Q function is maximum, which is

$$\theta^{k+1} = \arg\max Q(\theta \mid \theta^k) \tag{4}$$

2.3. Parameter estimation based on EM algorithm

For state-space model (1) and (2), let $p = (x_{1:T}, y_{1:T} \mid \theta)$, representing the likelihood functions of state and observation variables. The Q function was defined as the expectation of log-likelihood function $\log\left[p = (x_{1:T}, y_{1:T} \mid \theta) \right]$, shown in the following integral formula(Jing Deng 2013):

$$\begin{aligned} Q(\theta \mid \theta^k) &= E_{X \mid Y_0, \theta^k} \left\{ \log[p = (Y_0, X \mid \theta)] \right\} = \int \log[p = (Y_0, X \mid \theta)] p(X \mid Y_0, \theta^k) dX \\ &= \int \log[p = (x_{1:T}, y_{1:T} \mid \theta)] p(x_{1:T} \mid y_{1:T}, \theta^k) dx_{1:T} \end{aligned} \tag{5}$$

In Formula (5), the first part $p = (x_{1:T}, y_{1:T} \mid \theta)$ is the joint density function of state and output, which can be derived as follows:

$$p = (x_{1:T}, y_{1:T} \mid \theta) = p(x_1 \mid \theta) \prod_{t=2}^{T} p(x_t \mid x_{t-1}, \theta) \prod_{t=1}^{T} p(y_t \mid x_t, \theta) \tag{6}$$

Substitute Formula (6) into Formula (5) to get:

$$\begin{aligned} Q(\theta \mid \theta^k) &= \int \log[p(x_{1:T}, y_{1:T} \mid \theta)] p(x_{1:T} \mid y_{1:T}, \theta^k) dx_{1:T} \\ &= \int \log[p(x_1 \mid y_{1:T}, \theta)] p(x_{1:T} \mid y_{1:T}, \theta^k) dx_{1:T} \\ &+ \sum_{t=2}^{T} \int \log[p(x_t \mid x_{t-1}, \theta)] p(x_{1:T} \mid y_{1:T}, \theta^k) dx_{1:T} \\ &+ \sum_{t=1}^{T} \int \log[p(y_t \mid x_t, \theta)] p(x_{1:T} \mid y_{1:T}, \theta^k) dx_{1:T} \end{aligned} \tag{7}$$

To calculate the Q function in Formula (7), the value of density function $p(x_{1:T} \mid y_{1:T}, \theta^k)$ need to be obtained. Since it’s relatively complex to directly calculate the density function, particle filters were used to determine the value of the density function.

2.4. Estimation of density function through particle filter

Particle filter is to approximate probability density function $p(x_k \mid z_k)$ by looking for a set of random samples that spread in state space. Integral operation is replaced with sample mean to obtain the process of the minimum variance state estimation. And the samples are called "particles".

Based on the method in Literature [8] and [10], particle filters were utilized to estimate density function, computing the Q function. When the current parameter estimation θ^k was given, the density function of state can be approximated as:

$$p(x_t \mid y_{1:t}, \theta^k) \approx \sum_{t=1}^{N} \omega_t^i \delta(x_t - x_t^i) \tag{8}$$

where $\delta(.)$ is Dirac delta function, N is the number of particles, ω_t^i is the standardized weight of the ith particle, and $\sum_{i=1}^{N} \omega_t^i = 1$.

Consequently, the previous integral formula (3) based on posterior density function can be converted into a summation operation based on weighted samples. In many cases, the above posterior density can be multivariable, high-dimensional, multi-peak, non-standard and non-analytic. It's very difficult to directly sample particles from such density functions, so importance sampling (IS) (MS Arulampalam 2002; Chen Zhe 2003)was introduced. A posteriori can be obtained through re-sampling the importance sampling. When the number of samples tends to infinity, the estimation obtained from particle filters is asymptotically unbiased (S.D. Grantham and L.H. Ungar 1990). Generally, the probability distribution of state transition is taken as the importance density, such as:

$$q(x_t \mid y_{1:t}, \theta^k) = p(x_t \mid x_{t-1}, \theta^k) \tag{9}$$

At this point, the weight normalization for each particle can be expressed as follows:

$$\ddot{\omega}_t^i \propto \omega_{t-1}^i p(y_t \mid x_t^i, \theta^k) \tag{10}$$

$$\omega_t^i = \frac{\ddot{\omega}_t^i}{\sum_{i=1}^{N} \ddot{\omega}_t^i} \omega_{t-1}^i \tag{11}$$

However, the proposed algorithm has a major defect—the reduction of weights. The direct consequence is that the calculation of most particles about particle concentration has little significance due to their small weights. The most effective way to reduce this phenomenon is re-sampling. Its basic idea is to increase the number of particles with large weights by re-sampling particles and the probability density function expressed by corresponding weights. After re-sampling, the weight of each particle will be reset to $\omega_t^i = \dfrac{1}{N}$.

Although solving the reduction of particles, re-sampling reduce the diversity of the particles at the same time. So re-sampling is used only when necessary, rather than at each step. Therefore, the concept of effective sample size was introduced to measure the reduction of weights, defined as (A. X. Shao 2010):

$$N_{eff} = \frac{1}{\sum_{i=1}^{N} (\omega_t^i)^2} \tag{12}$$

where ω_t^i is the normalized importance weight. A threshold N_{th} can be designed in advance. Re-sampling is required when $N_{eff} \le N_{th}$, and the reduction of weights is relatively serious. Otherwise it is needless.

Combined with the current parameter estimation, the flow chart of the particle filter algorithm can be summarized as follows:

Step 1: Initialization.

Draw $\{x_0^i\}_{i=1}^N$ from $p(x_0 \mid \theta^k)$, set $\omega_0^i = \frac{1}{N}$,t=1.

Step 2: Importance sampling.

Generate $\{x_t^i\}_{i=1}^N$ from $p(x_t \mid x_{t-1}, \theta^k)$.

Step 3: Assigning weights.

Assign the weights $\{\omega_t^i\}_{i=1}^N$ using Eqs.(10)and(11), computer N_{eff} using Eqs.(12), if $N_{eff} \le N_{th}$, go to step 4,or go to step 5.

Step 4: Resampling.

Generate $\{x_t^i\}_{i=1}^N$, set $\omega_t^i = \frac{1}{N}$.

Step 5: set t=t+1, repeat step 2 to step 4 until t is more than T.

2.5. *EM algorithm based on particle filter*

The estimation for $p(x_{1:T} \mid y_{1:T}, \theta^k)$ is a smoothing problem for all state. Based on the iterative nature of EM algorithm, the smoothing for all state was calculated. Afterwards the method of Literature(A. Doucet 2001) was utilized to further marginalize the state. And the Q function can be further expressed in the following form:

$$
\begin{aligned}
Q(\theta \mid \theta^k) &= \int \log[p(x_1 \mid y_{1:T}, \theta)] p(x_1 \mid y_{1:T}, \theta^k) dx_1 \\
&+ \sum_{t=2}^{T} \int \log[p(x_t \mid x_{t-1}, \theta)] p(x_{t-1:t} \mid y_{1:T}, \theta^k) dx_{t-1:t} \\
&+ \sum_{t=1}^{T} \int \log[p(y_t \mid x_t, \theta)] p(x_t \mid y_{1:T}, \theta^k) dx_t
\end{aligned}
\tag{13}
$$

The calculation for $p(x_t \mid y_{1:T}, \theta^k)$ and $p(x_{t-1}, x_t \mid y_{1:T}, \theta^k)$ are also smoothing problems with a large amount of calculation. One available solution is: When t=1:T, replace $p(x_t \mid y_{1:T}, \theta^k)$ with $p(x_t \mid y_{1:t}, \theta^k)$. When t=1:T-1, replace $p(x_{t-1}, x_t \mid y_{1:T}, \theta^k)$ with $p(x_t, x_{t+1} \mid y_{1:t}, \theta^k)$ (J. Deng and B. Huang 2012). In practical applications, the particle filter algorithm can greatly reduce the complexity of calculation. The algorithm can improve the performance of the EM algorithm to some extent through iterative calculation.

In Formula (13), through particle filter, the density function $p(x_t \mid y_{1:T}, \theta^k)$ can thus be approximated as:

$$
p(x_t \mid y_{1:T}, \theta^k) \approx p(x_t \mid y_{1:t}, \theta^k) = \sum_{i=1}^{N} \omega_t^i \delta(x_t - x_t{}^i) \tag{14}
$$

The joint density function of x_t and x_{t+1} can be approximated as:

$$
\begin{aligned}
& p(x_t, x_{t+1} \mid y_{1:T}, \theta^k) \approx p(x_t, x_{t+1} \mid y_{1:t}, \theta^k) \\
& = p(x_{t+1} \mid x_t, \theta^k) p(x_t \mid y_{1:t}, \theta^k) \\
& = \sum_{i=1}^{N} \omega_{t|t+1}^i \delta(x_t - x_t{}^i) \delta(x_{t+1} - x_{t+1}{}^i)
\end{aligned}
\tag{15}
$$

where

$$
\omega_{t|t+1}^i = \frac{p(x_{t+1}^i \mid x_t^i, \theta^k) w_t^i}{\sum_{i=1}^{N} p(x_{t+1}^i \mid x_t^i, \theta^k) w_t^i} \tag{16}
$$

Using the approximation of these density functions, the Q function can be written as the following form:

$$Q(\theta \mid \theta^k) \approx \sum_{i=1}^{N} \omega_1^i \log[p(x_1^i \mid \theta)] + \sum_{t=2}^{T}\sum_{i=1}^{N} \omega_{t-1|t}^i \log[p(x_t \mid x_{t-1}^i, \theta)] + \sum_{t=2}^{T}\sum_{i=1}^{N} \omega_t^i \log[p(y_t \mid x_t^i, \theta)] \quad (17)$$

After obtaining the approximation of the Q function, EM algorithm can be used for calculation. In Step Expectation, the current estimated parameter θ^k was utilized to estimate the Q function. In Step Maximization, the latest parameter θ^{k+1} was obtained through maximizing the Q function. The parameter values when the Q function is maximum can be obtained by calculating the differentiation of the Q function about θ. In each iteration, the parameter estimation is optimal when $\partial Q / \partial \theta_j = 0$, where θ_j is the jth system parameters.

EM algorithm based on PF (PF-EM algorithm for short) can be summarized as follows:

Step 1: Initialization.

Parameter θ is initialized, and t=0.

Step 2: Expectation.

At the time of t, Formula (17) is used to calculate the estimation of the Q function, and the current system parameter θ^k is given.

Step 3: Maximization.

Through maximizing the proximate formula of the Q function, the latest parameter θ^{k+1} can be obtained. And let k=k+1.

Step 4: Repeat Step 2 and 3 until the convergence condition is satisfied. For example, the change of system parameters between two iterations is less than the tolerance.

3. Design for soft sensor of COx content in tail gas of PX oxidation side reactions

As an important raw material of polyester fiber and plastic,Purified terephthalic acid (PTA) can be obtained from PX through an oxidation action in AMOCO process, whose catalyst is Mn and Co, while promoter is Br, and acetic acid (HAc) is its solvent (Prengle Jr H W and Barond N 1970). There are a large number of side-reaction besides in which product is CO and CO_2 mainly(Chunyang Zhang 1999; Yun Ming 2002; Dehua wu 2000).Leading a drop in the quality of products, and an jump in the consumption of catalyst and promoter.

In this paper, the contents COx was studied using the proposed algorithm on the base of the plants in Yang Zi Petrochemistry Company (YPC).After principal component analysis, six variables were choosed as the input data, including reaction temperature(u_1 ， ℃ ） , solvent ratio(u_2 ， Kg.HAc/Kg.PX), Co concentration(u_3 ， wt%), Mn concentration(u_4 ， wt%), Br concentration(u_5 ， wt%)and reaction time(u_6 ， s)respectively. All industrial data presented in this paper have been normalized.

PX oxidation side reaction is a continuous process. And the CO_X content in its tail gas, related to past values, was described through dynamic model.The ultimate nonlinear state-space model was shown as follows which was proved to be feasible:

$$\begin{aligned} x_t &= a x_{t-1} + B^T u_{t-1} + c\cos(x_{t-1}) + \omega_t \\ y_t &= x_t + v_t \end{aligned} \quad (18)$$

where the system parameter to be identified $\theta = [a, b_1, b_2, b_3, b_4, b_5, b_6, c]^T$, x_t represents the state vector, y_t represents the output vector, and $u_t = [u_{1t}, u_{2t}, u_{3t}, u_{4t}, u_{5t}, u_{6t}]^T$ represents the input vector. The number of inputs in this model is 6. ω_t and v_t , as independent identically distributed gaussian noises, respectively represent process noise and observation noise. Their covariance are respectively Q=0.05 and R=0.05. The proposed EM algorithm based on particle filter was adopted to estimate the parameters of the state-space model. In the Step Expectation of EM algorithm, the Q function was calculated according to Formula (17), where

$$\log[p(x_t \mid x^i_{t-1},\theta) = \log[\frac{1}{\sqrt{2\pi Q}}\exp[-\frac{1}{2}\frac{(x^i_t - ax^i_{t-1} - Bu_{t-1} - A\cos(x^i_{t-1}))^2}{Q}]] \tag{19}$$

$$\log[p(y_t \mid x^i_t,\theta)] = \log[\frac{1}{\sqrt{2\pi R}}\exp[-\frac{1}{2}\frac{(y_t - x^i_t)^2}{R}]] \tag{20}$$

Through making the derivative of Q function equal 0, parameters of the current step $\theta_{new} = [a_{new}, b_{1new}, b_{2new}, b_{3new}, b_{4new}, b_{5new}, b_{6new}, c]^T$ can be determined via that of the previous step during iteration. Then the parameter estimation proposed in this paper was utilized to determine the system parameters. The number of particle filters is 100, and the initial system parameters were [0.5, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 1].

The proposed soft sensing method was used for simulation analysis. The trend comparison between the true value of state and the PF-EM soft-sensing estimation was shown in Figure 1. The change of squared error was shown in Figure 2. The relation between sample space and posterior density over time was shown in Figure 3.

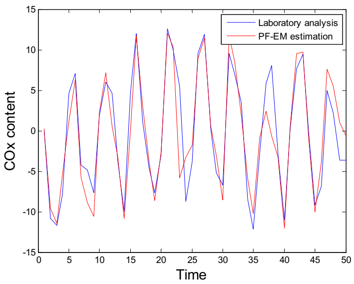

Fig.1. Trend comparison

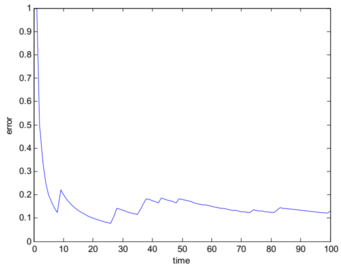

Fig.2 The change of squared error

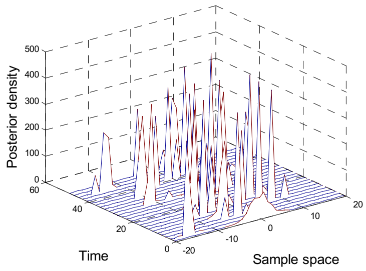

Fig.3.Trend comparison of sample space and posterior density

Figure 1-3 show that the proposed algorithm can effectively identify nonlinear model. And the predicted variables can well keep track of the change of the actual state. After finite iterations, the system error tends to 0. Re-sampling the particles sampled in the importance sampling functions via particle filters can excellently approximate the posterior density.

4. Discussion

An implementation method for soft sensor of nonlinear state-space models was proposed in the work. Thus the parameter estimation of nonlinear system was achieved based on particle filter and EM algorithm. In the soft sensing applications, most existing methods are achieved through deviation compensation based on the different renewal deviations between the prediction and practical observations. Therefore, the proposed algorithm has a great potential. However, there are still many issues worthy of further research to guarantee the reliability of practical industrial applications, such as delay in the test data, robustness of the prediction and mutation of the predicted values due to updating the experimental data. The influence of the time delay in laboratory sampling can be reduced in inline filtering. Robustness of the prediction can be improved via mechanism models. And mutation of the predicted values can be reduced by tuning of variance in state-space models. Thus the predictive quality of the model can be enhanced by combining these techniques.

5. Conclusions

Identification for nonlinear data-driven state-space model was proposed in the work. EM algorithm was used to estimate system parameters, while particle filters approximate expectation functions. The proposed PF-EM algorithm can identify the practical industrial processes due to small amount of calculation. Simulation results show that the proposed algorithm can effectively identify nonlinear system with an estimation error within the allowable range. Thus it can be extended to other industrial processes.

Acknowledgements

The authors would like to thank the financial support provided by Research and Innovation Project for College Graduates of Jiangsu Province of China under Grant CXLX12_0648, National Natural Science Foundation under Grant 61273142 and Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

References

[1] Brosillow. Inferential control of process[J]. Jour. AichE, 1978, 24(3): 485-509.
[2] P. Kadleca, B. Gabrys, S. Strandtb, Data-driven soft sensors in the process industry, Computers and Chemical Engineering :33 (2009) 795–814.
[3] S. Khatibisepehr, B. Huang, Dealing with the irregular data in soft sensors: Bayesian method and comparative study, Industrial and Engineering Chemistry Research :47 (2008) 8713–8723.
[4] W. Yan, H. Shao, X. Wang, Soft sensing modeling based on support vector machine and Bayesian model selection, Computers and Chemical Engineering:28 (2004) 1489–1498.
[5] S.D. Grantham, L.H. Ungar, A first principles approach to automation trouble shooting of chemical plants, Computers and Chemical Engineering: 14 (1990)783–798.
[6] P.K. Pearson, Control systems, identification, Encyclopedia of Physical Science and Technology (2004) 687–707.
[7] A. Doucet, N. de Freitas, N. Gordon, Sequential Monte Carlo Methods in Practice, Springer, New York, 2001.
[8] R.B. Gopaluni, A particle filter approach to identification of nonlinear process under missing observations, The Canadian Journal of Chemical Engineering :86(2008) 1081–1092.
[9] Jing Deng, Li Xie, Lei Chen, Shima Khatibisepehr, Biao Huang, Fangwei Xub, Aris Espejo. Development and industrial application of soft sensors with on-line Bayesian model updating strategy. Journal of Process Control: 23 (2013) 317– 325.
[10] J. Deng, B. Huang, Identification of nonlinear parameter varying systems with missing output data, AIChE Journal (2012), http://dx.doi.org/10.1002/aic.13735.
[11] MS Arulampalam.A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking [J].IEEE Trans.Signal Processing,2002,50(2):174-188.
[12]Chen Zhe. Bayesian Filtering: From Kalman Filters to Particle Filters, and beyond[R].Hamilton; McMaster University,2003.
[13] X. Shao, B. Huang, J. Lee, F. Xu, A. Espejo, Constrained Bayesian state estimation-a comparative study and a new particle filter based approach, Journal of Process Control: 20 (2010) 143–157.
[14] Prengle Jr H W,Barond N. Make Petrochemicals by Liquid Phase Oxidation [J]. Hydrocarbon Process, 1970a, 49(3):106.
[15] Prengle Jr H W,Barond N. Make Petrochemicals by Liquid Phase Oxidation(2).Kinetics,Mass Transfer and Reactor Design [J]. Hydrocarbon Process,1970b,49(11):150.
[16] Chunyang Zhang. Optimization of the oxidation process parameters of PTA reduces the consumption of

acetic acid[J]. Polyester industry,1999,12(3):39-42.
[17] Yun Ming. Analysis of reducing PX consumption in PTA production[J]. Polyester industry,2002,15(5):35-37.
[18] Dehua Wu, Chao Huang,Wenge Wu. Optimization of TA technology to reduce the consumption of acetic acid[J]. Polyester industry,2000,13(2):37-40.